Deep Generative models for 3D representations

Three-dimensional data, like point clouds and RGB-D, are becoming increasingly popular. They are being applied in modern technology, wherever the spatial computer vision is used, such as drones or autonomous vehicles. Those applications are usually used in scenarios that employ real-time decision making, forcing the design of the algorithms that process the data with both the accuracy and the efficiency in mind.

One of the most commonly used forms of three-dimensional data is point clouds. Point clouds are obtained with the usage of the devices such as LIDAR, which are registering the surroundings by laser and map them as points in three-dimensional space. Importantly, such a representation of data does not enforce any structure of the data (unlike in, e.g. image data, where every pixel has its own, constant position). Such freedom allows for an accurate projection of the LIDAR's environment, but on the other hand, it complicates significantly the procedures that operate on the collected data. As an example, one can imagine the situation, where a given point cloud is represented twice, but in both cases, the points are saved in different orders. It is important, that the employed approach was able to recognize those both point clouds as the same object, what in turn, can save from the excessive complexity and long execution time of the algorithms. Another difficulty related to 3d point clouds is the size of the data. During a single scan, the LIDAR device can create a representation of the surroundings that consists of a few thousand and even up to several dozen million points. This is another challenge for designing efficient ways of processing 3D data.

One of the problems in the are of machine learning is the problem of object generation. This problem involves building a model that will be able to construct previously unseen objects, that are consistent with underlying data distribution. What it means in practice is that the generative model trained on the images of faces will be able to generate the faces of nonexisting people. Currently, there exist several methods of object generation that work on image data. Interesting extensions of those methods are solutions applied to 3D objects, which are the topic of this project. Generating 3D objects may be useful in spatial design, including autonomous vehicles or augmenting the training dataset. Furthermore, the generative models may be used for constructing a set of features (also called a *latent representation*) with significantly lower dimensionality, but still preserving as much as possible of the information about the original data.

The goal of this project is the construction of several generative methods for 3D representations. Those methods are expected to be trained in an effective manner, generate high quality of the 3D objects samples and operate on real-valued representations. Additionally, presented methods will be extended with an ability to find binary representations, i.e. consisting only of values 0 and 1, which further improves the processing speed (on the hardware level the speed of processing the integer numbers are usually 1-2 orders of magnitude faster than floating-point numbers). Developed methods will be analyzed qualitatively and quantitatively using the benchmark datasets publicly available on the Internet.